

Progress Report for NASA AISR Award NNG06GE96G “Novel Higher-Order Statistical Method for Extracting Dependencies in Multivariate Geospace Data Sets (September 2008-October 2009)”

Principle Investigator: Jay R. Johnson
Co-Investigator: Simon Wing

Most Important Results for this Year

In the first year of this project, we examined time series of variables that characterize the magnetospheric state and how they depend nonlinearly on the solar wind driver. This year, we focused on practical aspects of how to apply this information to construct predictive models of magnetospheric dynamics. In particular, we studied how statistical measures of information content could be used to improve performance of neural network models of the magnetosphere.

One area that should be a concern for modeling is the issue of noise in data sets. Noise can be a major issue when trying to train neural network models as illustrated in Figure 1. In Figure 1a, we have taken a known nonlinear system (Mackey-Glass) and added noise to the system. We have trained a neural network for the system using a vector of four points in the time series $[x(t), x(t-\Delta), x(t-2\Delta), x(t-3\Delta)]$ (shown with green dots) to predict $x(t+p)$ (shown in red dots). With the addition of noise, the underlying dynamics is poorly captured and the performance of the neural network is not adequate. On the other hand, when cumulant-based significance is used to identify noisy data, that data can be eliminated from the training set. In this case, although the neural network was trained using data selected from the same time noisy time series, the underlying dynamics is well-captured because the noisy data has been eliminated.

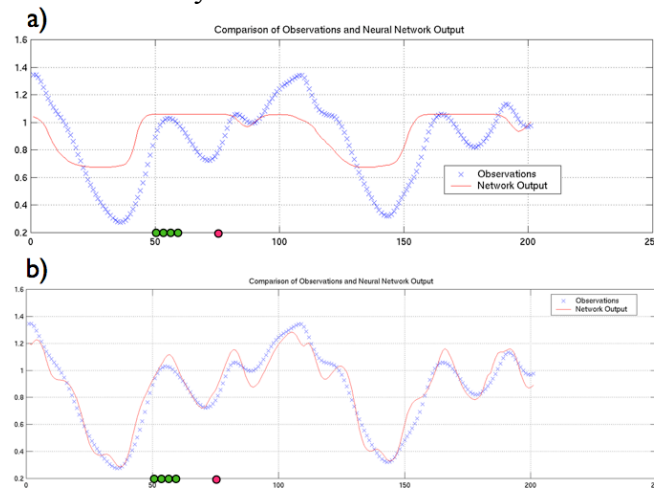


Figure 1. A neural network is trained to predict the evolution of a Mackey-Glass time series given four prior values (i.e. green dot values are used to predict red dot value). Panel (a) shows poor performance when sporadic noise is added to the original time series. The original time series is shown with blue “x” and the predictive neural network model is shown in red. Panel (b) shows significantly improved performance when

cumulant-based significance is used to identify noisy data and eliminate it from the training set.

Cumulant-based measures of significance are also a good indicator of changes in underlying dynamics of the magnetosphere and indicate a dynamical transition in the system. Therefore, changes in the underlying dynamics can be identified by looking for changes in the cumulant-based significance, which indicate nonstationarity. We have shown that such changes in dynamics occur over the course of a solar cycle with strong nonlinear dependence in the declining phase of the solar cycle just prior to solar minimum. In addition, there are also shorter scale variations in the dynamics. In Figure 2 we shown the significance of K_p as a function of time during 1987. It is apparent that there are periods where the cumulant-based significance increases for prolonged periods of time and also periods where it is smaller. This variation suggests changes in the dynamical state of the system that can be better modeled with neural networks that are trained specifically for high and low significance data. The cumulant-based significance can be used to sort the data for training multiple neural networks and also can be used as an indicator for state transition between networks.

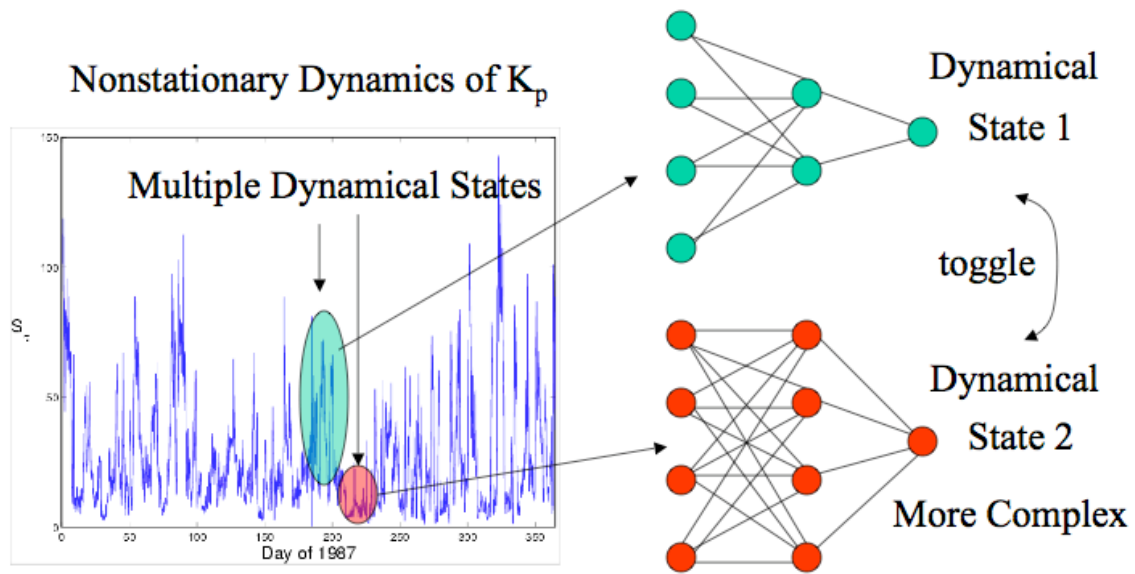


Figure 2. Nonstationarity in the dynamics of the magnetosphere during 1987 is evident in the variation of the cumulant-based significance which shows periods of high significance and low significance. This statistical measure can be used to select data based on the range of underlying dynamics to train networks and to identify transition probabilities between the networks.

Finally, an important success of our project this year has been the transitioning of improved neural network models of K_p and Dst (measures of the global state of the magnetosphere) to AFWA (Air Force Weather Agency), NOAA Space Weather Prediction Center (SWPC) and Air Force Research Laboratory (AFRL). We have developed 3 K_p forecast models:

- APL model 1: inputs ACE solar wind data and nowcast Kp and predicts Kp one hour ahead
- APL model 2: inputs ACE solar wind data and nowcast Kp and predicts Kp four hour ahead
- APL model 3: inputs ACE solar wind data and predicts Kp one hour ahead.

Our extensive evaluations show that all of our models perform better than the existing models. These are shown in Figure 3.

Similarly, we have developed 2 Dst models:

- APL model 1: inputs ACE and predicts Dst one hour ahead.
- APL model 2: inputs ACE and predicts Dst four hour ahead.

These two Dst models have been transitioned to AFWA.

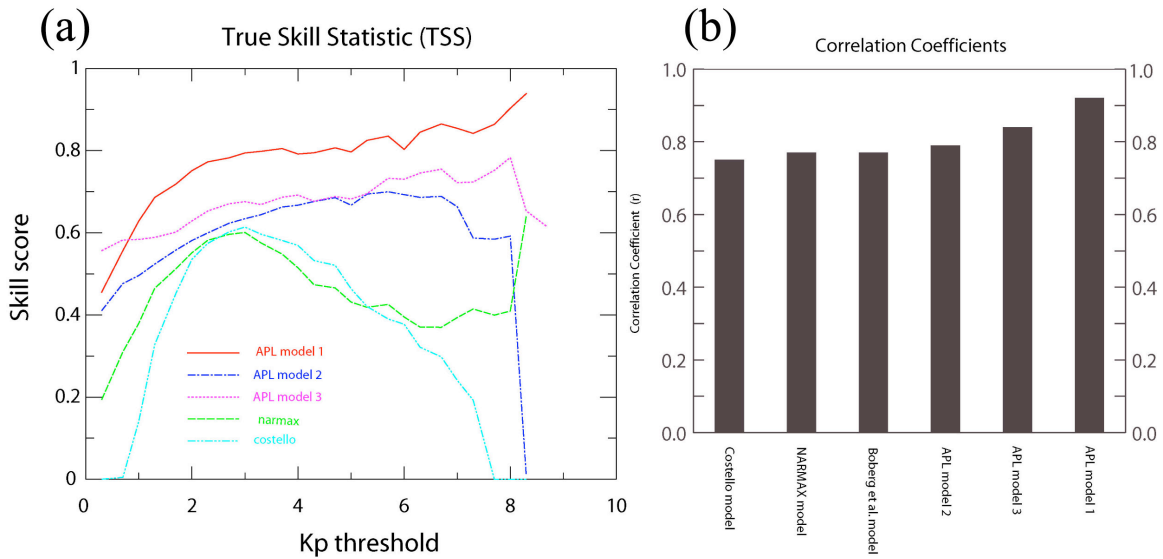


Figure 3. Skill scores (a) and correlation coefficients (b) show that APL Kp forecast models perform better than other Kp forecast models. The evaluations were done using IMP8, Wind, and ACE solar wind data spanning the period 1975 – 2001.

Summary Of Progress

- Examined how noise affects entropy-based discriminating measures of nonlinear dependency. We found that typically, moment-based cumulant measures can give better statistics in high-dimensional systems where it is more difficult to construct the probability distribution and construct a mutual information measure.
- Used cumulant-based significance to identify noisy data for the purpose of improving neural network models of space plasmas

- Identified nonstationarity in magnetospheric dynamics using cumulant-based significance, and have developed a scheme to use it to select data for training of specialized neural networks and to indicate transition probabilities between the networks. We are in the process of testing the method on known systems.
- Transitioned neural network models to AFWA, NOAA SWPC, and AFRL for widespread usage.
- Tested transfer entropy as a measure of causal relationships in space plasmas with applications to solar-wind/magnetosphere coupling showing improvements over studies based on cross-correlations.
- Applied transfer entropy and mutual information to substorm data sets to identify causal relation between external and internal triggers of substorms. Our preliminary results indicate that additional information about external triggers (such as northward IMF turning) do not add additional information suggestive that northward IMF turning may not be causally related to substorm onset.

Presentations

- Johnson, J. R., and S. Wing, An information-theoretical approach to identify nonlinearity in magnetospheric activity, 2007 Joint Assembly, 22-25 May 2007, Acapulco, Mexico.
- Wing, S., and J. R. Johnson, Kp forecast models, IUGG XXIV General Assembly, July 2-13, 2007, Perugia, Italy.
- J. R. Johnson and Wing, S., and J. R. Johnson, Variation of Magnetospheric Dynamics During the Solar Cycle, IUGG XXIV General Assembly, July 2-13, 2007, Perugia, Italy.
- Simon Wing and J. R. Johnson, "Forecasting geomagnetic activity indices," Fall AGU meeting, 2007.
- Jay R. Johnson and Simon Wing organized a special session on "Entropy constraints on space plasmas for transport and phase transitions," Fall AGU meeting, 2007.
- Jay R. Johnson attended the NASA AISRP PI meeting and presented an invited talk focusing on our recent efforts at detecting nonlinearity in space plasmas and application of transfer entropy to understand causal relationships in space plasmas.
- Jay R. Johnson and Simon Wing "Detecting Causality In Space Plasmas With Entropy Based Measures" presented at Fall AGU meeting, 2008.
- Jay R. Johnson and Simon Wing, "Detecting Causality In Magnetospheric Dynamics using Transfer Entropy" at the International Association of Geomagnetism and Aeronomy, Aug 23-30 2009.
- Simon Wing and Jay R. Johnson, "Predicting Geomagnetic Activity Indices" at the International Association of Geomagnetism and Aeronomy, Aug 23-30 2009.
- Simon Wing attended NASA AISRP PI meeting (Oct 14-16, 2009) and presented an invited talk on our NASA project emphasizing our recent efforts to broadly disseminate our predictive models for geomagnetic indices through transfer to NOAA.
- Jay R. Johnson and Simon Wing served as guest editors of a special section in the Journal of Geophysical Research that focused on application of entropy based

measures as constraints on magnetospheric dynamics including the development of entropy-based tools to study nonlinearity, *vol. 114, no. A9, 2009*.

- Gave presentations at Augsburg college, West Virginia University, and Naval Research Laboratory about nonlinearity in the magnetosphere.

Papers:

- Zhu, D., S. A. Billings, M. A. Balikhin, S. Wing, and H. Alleyne (2007), Multi-input data derived Dst model, *J. Geophys. Res.*, 112, A06205, doi:10.1029/2006JA012079.
- Wing, S., and Jay R. Johnson, Introduction to special section on entropy properties and constraints in the space plasma, *J. Geophys. Res.*, doi:10.1029/2009JA014911, in press.
- Wing, S., and J. R. Johnson Substorm entropies, *J. Geophys. Res.*, 114, A00D07, doi:10.1029/2008JA013989, 2009.
- Johnson, J. R., and S. Wing, Northward interplanetary magnetic field plasma sheet entropies, *J. Geophys. Res.*, 114, A00D08, doi:10.1029/2008JA014017, 2009.
- Jay R. Johnson and Simon Wing, "Detecting Causality In Space Plasmas With Entropy Based Measures," to be submitted to *J. Geophys. Res.*, 2009.
- Jay R. Johnson and Simon Wing, "Nonlinearity in Kp and High-Speed Stream Interfaces," currently under preparation for publication in *J. Geophys. Res.*, 2009.

Tasks To Be Carried Out In The Coming Year

In the coming year, we will:

- Identify good state variables for the magnetosphere based on content of information about the nonlinear response of the magnetosphere to solar wind. We will test tail stretching and auroral activity index such as AE. We will consider causal relationships between the state variables using transfer entropy.
- Identify a coupling function based using maximization of information flow between the solar wind and magnetospheric variables. We have compiled the appropriate data sets and have already developed tools to perform this task based on maximization of entropy. This coupling function should be an improvement over ad hoc coupling functions recently suggested in the literature.
- Analyze a database of substorm events using transfer entropy as a discriminating statistic to examine causality in substorms to determine substorm triggers comparing internal vs external mechanisms.
- The solar wind – magnetospheric state variables coupling functions identified in the above analyses should help improve the model predictions for Kp, Dst, magnetotail stretching and AE. For example, the coupling functions could be explicitly used as input parameters in the forecast models. If time and funding permits, we will develop models or improve the existing models to predict these magnetospheric state variables.